DOCUMENT RESUME

ED 423 310 TM 029 198

AUTHOR	Rudner, Lawrence
TITLE	Item Banking. ERIC/AE Digest.
INSTITUTION	ERIC Clearinghouse on Assessment and Evaluation, Washington, DC.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	EDO-TM-98-04
PUB DATE	1998-08-00
NOTE	4p.
CONTRACT	RR93002002
AVAILABLE FROM	ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, Schriver Hall, College Park, MD 20742-5701; telephone: 800-464-3742 (free).
PUB TYPE	ERIC Publications (071) -- ERIC Digests in Full Text (073)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Adaptive Testing; *Computer Assisted Testing; Difficulty Level; *Item Banks; Item Response Theory; *Test Construction; Test Items
IDENTIFIERS	ERIC Digests; *Rasch Model

ABSTRACT

This digest discusses the advantages and disadvantages of using item banks, and it provides useful information for those who are considering implementing an item banking project in their school districts. The primary advantage of item banking is in test development. Using an item response theory method, such as the Rasch model, items from multiple tests are placed on a common scale, one scale per subject area. The scale indicates the relative difficulty of the items, and items can be placed into the scale without extensive testing. New subtest and tests with predictable characteristics can be developed by drawing items from the bank. Another advantage of the item bank is that the test developer can "deposit" or "withdraw" items as needed. Large deposits can be made by merging the item bank with one from another district, and small deposits can come by adding only a few items. Another advantage to item banking is that it helps establish a language for discussing curriculum goals and objectives. However, item banking and item response theory are not cure-alls for measurement problems. Care and effort must still go into item writing. A great deal of work must go into preparation and planning. Computer experts should be available to: modify computer programs; establish a database system; and run packaged programs. The most crucial step is planning, which involves preparing and training those who will work with the item bank, the identification of the initial contents, and the identification of what the developer hopes to accomplish with the item bank. The start-up activities would mostly involve administrative activities and the data processing staff. Running the item bank then involves depositing new items, with field testing as necessary, and ensuring that the bank is used appropriately. As a service to instructional and curriculum staff, item bank developers can provide information on the relative difficulty of different tasks within and across grade levels. (Contains five references.) (SLD).

ED 423 310

Item Banking

Lawrence Rudner
ERIC Clearinghouse on Assessment and Evaluation

TM029198

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ERIC®/AE Digest Series EDO-TM-98-04 **August 1998**
The Catholic University of America **Department of Education**

ERIC®/AE Digest Series EDO-TM-98-04 **August 1998**
The Catholic University of America **Department of Education**

Item Banking

Lawrence Rudner, The ERIC Clearinghouse on Assessment and Evaluation

Introduction

Various school districts use standardized tests as a way to measure scholastic achievement. Usually, these districts need to revise tests with some frequency to avoid administering the same test year after year. Unfortunately, creating new tests can be a very time consuming endeavor. Not only do test writers need to compose the test items, they also must determine each item's difficulty in order to ensure that a test will neither be too hard nor too easy.

Using item banks, test makers can escape this process. Item banks are files of various suitable test items that are "coded by subject area, instructional level, instructional objective measured, and various pertinent item characteristics (e.g., item difficulty and discriminating power)" (Gronlund, 1998, p. 130). The purpose of this digest is to discuss the advantages and disadvantages of using item banks as well as provide useful information to those who are considering implementing an item banking project in their school district.

Advantages of Item Banking

The primary advantage of item banking is in test development. Using a item response theory method, such as the Rasch model, items from multiple tests are placed on a common scale, one scale per subject matter. The scale indicates the relative difficulty of the items. Items can be placed on the scale, i.e. into the item bank, without extensive testing. New subtests and tests, with predictable characteristics, can be developed by drawing items from the bank. For example, suppose you are interested in developing a new subtest to cover fractions in seventh grade. You can go to the item bank, identify items related to your objectives and then predict the characteristics of a subtest composed of those items. The effect of including or excluding particular items can also be predicted.

Another advantage of an item bank is that it will permit you to "deposit" additional items to be withdrawn as needed. Depending on the size of the testing program, there can be two practical approaches for making deposits. You can make "large deposits" by merging your item bank with one from another district. You can also make "small deposits" by adding a few locally developed items at a time. The large deposit option will involve purchasing or trading items with another district and then equating their scale to yours. The small deposit option involves piloting a fewer number of items with examinees in several grade levels. This can easily be accomplished by adding a supplemental page containing experimental items to be administered along booklet from the school system.

Item banking provides substantial savings of time and energy over conventional test development. In traditional test development, items can only be described relative to the other items within the test and to whom they were given. That is, item characteristics are extremely group and test specific. With item banking, items are described their relative difficulty across grade levels. In order to develop a new test or subtest, one does not need to go through the laborious process of developing a large set of items for piloting and evaluating. Instead, one just draws from the bank. Further, drawing from the bank allows one to make fairly accurate predictions concerning composite test characteristics.

One additional advantage of item banking is that it helps establish a language for discussing curriculum goals and objectives. The items describe individual tasks students are capable or incapable of doing. The location of the items on a calibrated scale allows one to identify the relative difficulty of particular tasks. This provides a way to discuss possible learning hierarchies and ways to better structure curriculum.

Disadvantages and Limitations of Item Banking

Item banking and item response theory are not cure-alls for measurement problems. Persistence and good judgement must remain vital aspects in any test construction and test usage effort. One must make every possible effort to include only quality items in the item bank. The same care and effort must go into item writing. Items purchased form external sources must be evaluated carefully for match to your curriculum as well as for technical quality.

Item banking involves equating various tests and items. It is entirely possible, mathematically, to equate tests which cover entirely different subject matter. At the practical level, this means that it is also possible to equate items which assess subtly, but significantly different skills. In order to avoid this undesirable situation, the item review process must also include a careful evaluation of the skills assessed by each item and tests must be carefully formulated.

The intent of compiling a test using latent trait theory is to be able to make a prediction of the composite test characteristics. While the prediction is often surprisingly accurate, it must be validated. Tests developed using latent trait theory should still be field tested.

While some districts have implemented very successful item banks and Rasch calibrated testing programs without knowing anything about IRT, good practice calls for a staff that is comfortable with and knowledgeable of what they are doing. A district undertaking an item banking project should have full understanding of the practical as well as the mathematical/ theoretical aspects of item banking.

An item bank really consists of multiple collections of items with fairly unidimensional content area, such as mathematic computations or vocabulary. Collections of items usually span several grade levels. In order to develop the bank, many tests must be calibrated, linked (or equated), and organized. This requires a great deal of work in terms of preparation and planning and in terms of computer time and expertise. Once the item bank is established, however, test development time, effort, and cost is reduced.

Planning for an Item Bank

The most crucial step in developing an item bank is planning. This involves the preparation of individuals, the identification of what you have to start an item bank, and the identification of what you hope to accomplish with an item bank.

Everyone on the staff should have enough familiarity with Rasch measurement principles and item banking to be able to knowledgeably discuss and explain the project. You can formally train your staff by using in-house personnel, bringing in a traveling

ERIC® ERIC Clearinghouse on Assessment and Evaluation, 210 O'Boyle Hall, The Catholic University of America, Washington, DC 20064 • 800 464-3742 CUA

workshop, or having people attend a pre-session at a research association or conference.

You should have senior level personnel available to answer technical questions that might arise. You should also have computer experts that are capable of doing the following tasks: 1.) modifying computer programs, 2.) establishing a data base system, and 3.) capable of running packaged programs.

If you intend to do any item bank exchanges or purchases, you should have someone on your staff who knows what is available. You need personnel capable of critically evaluating test items for technical quality, curriculum match, unidimensionality, and potential bias. In order to accurately calibrate test items and establish scales, items need to be presented to examinees with a wide range of ability.

In order to link various forms and grade levels within a content area, common anchor items are needed. (These anchor items must be administered along with the items within a given form. The form and anchor items are calibrated together. The anchor item parameter values based on calibration with one form are compared with the anchor item parameter values based on calibration with another form. The difference in parameter values is used to link the forms.) You need to identify for which content areas you have administered overlapping subtests and the number of students responding to the set of items. You may find you will need to gather additional item response data to link forms and grade levels.

Your data processing staff should examine literature and programs on item banking to determine what programs must be developed and what programs can be modified.

As much as possible, you should identify your projected testing needs for the next five years. This would involve identification of which subtests you will need to revise, what additional areas you may need to assess, and how objectives might be differently stressed.

Start-up Activities

The start-up activities would mostly involve administrative activities and the data processing staff. Each test would have to be calibrated and equated to the parallel form and adjacent grade levels. The data processing staff would have to adapt existing computer programs to the local system and develop a database system. They would then calibrate each test, equate the tests, and store the equated item parameters and their descriptors in a database system. With a large number of tests and items, this becomes a major undertaking.

Administrative staff would have to coordinate activities to insure that the data requirements are met. During the planning process, a chart can be developed to identify which tests and anchor items have been and will need to be administered to the requisite sample. Working from these charts, testing coordinators will need to organize the administration of tests and subtests needed to calibrate and equate all the items going into the item bank. This involves compiling test booklets, making testing arrangements, collecting response sheets, and preparing data for data processing. Depending on frequency of students taking multiple subtests from different levels and forms, this too can be a major undertaking.

Running the Item Bank

The item bank will allow you to withdraw items as needed to develop new or even special tests and subtests. There are basically two activities involved in running an item bank – making deposits and withdrawing items to develop a test.

As mentioned earlier, there are to viable options for making deposits to the item bank. The "large deposit" option involves merging an existing item bank with your own. If the existing item bank has been IRT calibrated, then you only need to administer a subset of items (per content area) from the new bank along with items already in your item bank. Remember, each item bank uses its own anchor items and allows you to equate the scales. This part will involve testing with a relatively small group of students. The anchor items from the new item bank can be appended to present group. Coordination would be similar to that involved in starting your own item bank.

The major task involved in using items from another item bank is a thorough, careful review of the items. All potential entries must be evaluated for technical quality, curriculum match, and potential bias. This would involve your test development experts, curriculum/instructional staff, and coordination between the two.

After an item review, items from non-calibrated could be treated like items developed by your staff. "Small deposits" would be made by calibrating and equating a few items at a time. One very efficient approach to collecting the requisite data is to append subtests of new items to original groups. The items within the original group would serve as anchor items for the new subtest(s) of items. In this manner, you can be constantly adding to your item bank.

Once developed and growing, your item bank is ready to provide the advantages discussed above. To develop a new subtest, you would develop a blueprint/table of specifications to outline what you want your new subtest to be like. Curriculum specialists and test development experts would then go to the item bank and identify which item in the bank appear appropriate in terms of content and in terms of their relative difficulty. If they find an insufficient number of items, them can make arrangements to add new items to the bank.

If the bank contains a sufficient number of items of the appropriate nature, the items can be grouped to form a new subtest. Without pilot testing, the characteristics of this new subtest can be predicted. With reasonable accuracy, you will know how much skill an examinee needs to obtain any given total raw score on the new subtest. The prediction should be validated by administering the subtest to students having received appropriate instruction and students not having received such instruction. This can also be accomplished by appending items to the existing forms. This validation would need a sample as large as you used in field testing the original group.

An item bank provides a scale of relative difficulty of tasks that covers multiple grade levels and skills within content areas. As a service to the instructional/curriculum staff, you can provide information on the relative difficulty of different taks within and across grades levels. For example, you can identify which fraction problems seventh graders find as difficult as certain decimal problems; or you can identify which reading skills taught in fourth grade can be mastered by students in their grade. It could also be used to help organize special programs for gifted and remedial students.

Additional Reading

Grolund, N.E. (1998). *Assessment of Student Achievement. Sixth Edition.* Needham Heights, MA: Allyn and Bacon.

Lord, F.M. (1980). Applications of item response theory to practical testing problems. Hillsdale, N.J. : L. Erlbaum Associates.

Mengel, Bill E.; Schorr, Larry L. (1992) Developing Item Bank Based Achievement Tests and Curriculum-Based Measures: Lessons Learned Enroute. (ERIC Document Reproduction Number ED344915).

Ward, A.W.; Murray-Ward, M. (1994). Guidelines for the development of item banks. An NCME instructional module. *Educational Measurement: Issues and Practice, 13*(1), 34-39.

Wright, B.D.; Stone, M.H. (1979). *Best Test Design. Rasch Measurement.* Chicago, IL: MESA Press.

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract RR93002002. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education. **Permission is granted to copy and distribute this ERIC/AE Digest.**